

1

smaller range. The character with the highest confidence level may be selected as the recognized character.

Some computer software for object recognition uses parameters to allow the software to be adjusted. The use of parameters allows the software to be tuned in a laboratory to particular conditions simulating the environment of anticipated operation of the software.

Before the software is shipped as part of a product, the parameters are fixed at a constant level that yielded the optimum recognition in the laboratory simulation for that product.

For example, if a scanned image represents the image using pixels, each having a greyscale value of 0-255, one parameter of the optical character recognition software may be to identify which values correspond to a part of the image to be recognized, in order to distinguish that part of the document from the greyscale value of the background.

For example, a document received via a fax that is photocopied onto off-white paper may have text that has a greyscale reading of 200, while the remainder of the page may have a greyscale reading of 100. A printed black and white document may have a greyscale reading of 240 for text and 30 for the remainder of the page. Text on a printed

002080"080700
0963473-080700

color document may have a greyscale reading as low as 90
with a greyscale reading of 70 for portions of the
background. These various values may be used to determine
that an optimal cutoff greyscale reading of 150 should be
5 used for the software. While this value provides a good
compromise for high-contrast documents such as most black
and white documents, certain color text on color background
documents simply will not be recognized with this parameter
value. If the parameter were lowered to 80 to accommodate
10 recognition of color documents, some black and white
documents would not be recognized, such as the fax
photocopied onto off white paper.

It would be desirable to have the parameter selection
process vary for each set of objects, such as characters on
15 the page, rather than selecting a single value for each
parameter and using that same value for all objects. This
would allow the parameter values to change for every page
or part of a page, causing the parameters to be optimized
for every circumstance. In the example above, it would be
20 desirable to use a greyscale threshold of 150 for the faxed
document and a threshold of 80 for the color document,
instead of using a value of 150 every time.

While it is possible to make several attempts at recognizing the objects, such as characters in the file, using different parameters for each attempt, and then selecting the attempt that yields the highest recognition confidence, such a process would add too much time to the recognition process to be practical. Although computing power increases every year, because users prefer to use the additional computing power to process images of higher resolution rather than improve the accuracy of the recognition, making several attempts at recognizing an image could take too long to be useful.

What is needed is a method and apparatus that can optimally set the parameters of an optical recognition without significantly adding time to the recognition.

Summary of Invention

A method and apparatus receives a digitized representation of an image at a first resolution, then produces a lower resolution version of the image. The method and apparatus then attempts multiple recognitions on the reduced-resolution version of the image using different parameters for each recognition and records the confidence level of the attempt. The parameters that yielded the highest confidence level on the attempt using the lower

resolution version can then be used to perform the
recognition on the higher-resolution version of the image.

Brief Description of the Drawings

Figure 1 is a block schematic diagram of a
5 conventional computer system.

Figure 2 is a block schematic diagram of a system for
identifying optimal parameters for optical object
recognition according to one embodiment of the present
invention.

10 Figure 3 is a flowchart illustrating a method of
identifying optimal parameters for optical object
recognition according to one embodiment of the present
invention.

Detailed Description of a Preferred Embodiment

15 The present invention may be implemented as computer
software on a conventional computer system. Referring now
to Figure 1, a conventional computer system 150 for
practicing the present invention is shown. Processor 160
retrieves and executes software instructions stored in
20 storage 162 such as memory, which may be Random Access
Memory (RAM) and may control other components to perform
the present invention. Storage 162 may be used to store
program instructions or data or both. Storage 164, such as

a computer disk drive or other nonvolatile storage, may provide storage of data or program instructions. In one embodiment, storage 164 provides longer term storage of instructions and data, with storage 162 providing storage for data or instructions that may only be required for a shorter time than that of storage 164. Input device 166 such as a computer keyboard, mouse, scanner, digital camera or any or all of these allows user input to the system 150. Output 168, such as a display or printer, allows the system to provide information such as instructions, data or other information to the user of the system 150. Storage input device 170 such as a conventional floppy disk drive or CD-ROM drive accepts via input 172 computer program products 174 such as a conventional floppy disk or CD-ROM or other nonvolatile storage media that may be used to transport computer instructions or data to the system 150. Computer program product 174 has encoded thereon computer readable program code devices 176, such as magnetic charges in the case of a floppy disk or optical encodings in the case of a CD-ROM which are encoded as program instructions, data or both to configure the computer system 150 to operate as described below.

In one embodiment, each computer system 150 is a conventional Intel Pentium/Celeron compatible computer such

as the model 6540C commercially available from Hewlett
Packard Corporation of Palo Alto, California, running the
Windows operating system commercially available from
Microsoft Corporation of Redmond Washington, although other
5 systems may be used.

Referring now to Figure 2 a system for identifying
optimal parameters for optical object recognition is shown
according to one embodiment of the present invention. The
description below uses characters as the objects, however,
10 objects can be shapes, people, three-dimensional items or
any other object. Native resolution image storage 210
receives at input 208 a digitized representation of an
image from an optical scanning device such as a scanner or
digital camera at input 208. Native resolution image
15 storage 210 contains conventional memory such as random
access memory or other types of storage and stores the
digitized representation of the image in this storage. As
native resolution image storage 210 receives an image, it
signals downsampler 212.

20 In one embodiment, the digitized representation of the
image stored in native resolution image storage 210 is made
up of an array of pixels at a resolution. For example, the
digitized representation of the image may be represented

resolution and may optionally compress the result prior to storage into low resolution image storage 214.

For example, if the resolution is 1200 dpi and it is desired to reduce the resolution to 300 dpi, downsampler 212 selects from native resolution image storage 210 a set of sixteen pixels from the upper left hand corner of the image, four pixel columns wide and four pixel rows tall, and averages the values from these pixels to produce a single pixel, which is output to low resolution image storage 214. Low resolution image storage 214 contains conventional storage such as memory or disk storage. The four pixels adjacent to these in each of the same four rows are selected and the process is repeated by downsampler 212 and so on until the end of the row is reached. When the end of the row is reached, downsampler 212 selects the next four rows and repeats the process described above for these rows. It isn't necessary to proceed step-by-step in this fashion, nor is it necessary to average the values. For example, the values may be smoothed using conventional smoothing techniques. When downsampler 212 has completed reducing the resolution of the digitized representation of the image, downsampler 212 signals parameter selector 220.

004080" E4EE960

Parameter selector 220 selects an initial set of one or more parameters and passes them to recognition engine 216. Recognition engine is any conventional recognition engine such as an optical character recognition engine that can accept a set of parameters, identify a segment of an image, and provide the one or more characters recognized at the highest confidence level and also supply that confidence level, which may be a value between 0 and 1.

Recognition engine 216 performs conventional optical character recognition techniques using the parameters provided by parameter selector 220 on the reduced resolution version of the digitized representation of the image that is stored in low resolution image storage 214. In one embodiment, recognition engine 216 segments the image into characters or words and then attempts to recognize the character or word using conventional optical character resolution techniques on each segment. Optical character recognition is described in Bunke & Wang, ed., *Handbook of Character Recognition and Document Image Analysis* (1997 World Scientific Publishing Co. Pte. Ltd, Singapore, ISBN 981-02-2270-X). Other forms of object recognition may also be used such as handwriting recognition, described in S Impedoio and J Simon, eds. *From Pixels to Features III*" (1992 Elsevier Science Publishers,

B.V., Amsterdam). Source code for an optical character recognition engine is publicly available from the National Institute of Standards and Technology, Gaithersburg, Maryland, 20899 and is described in Garriss et al, "Public Domain Optical Character Recognition", *Proceedings, SPIE* Volume 2422, pp2-15, and this engine may be suitably modified as described herein.

Parameter selector 220 selects another set of parameters and passes them to recognition engine 216 and signals recognition engine to repeat the recognition process described above on the same segment of the image it just processed, and recognition engine 216 complies with the request. This process may be repeated a number of times, each time with a different set of parameters.

In one embodiment, each time another set of parameters is selected and provided by parameter selector 220, the value of only one of the parameters is varied until several different values of that parameter have been selected and provided to recognition engine 216 by parameter selector 220. When all possible values of that parameter have been provided or when an acceptable value has been identified, that parameter is set to a particular value and the value

of a different parameter is varied, and so forth until all the parameters have been varied.

For example, assume two parameters, parameter 1 and parameter 2, with possible values of A1, B1, C1 and D1 for parameter 1 and values A2 and B2 for parameter 2. The sets of parameters provided by parameter selector 220 could be (A1, A2), (B1, A2), (C1, A2), (D1, A2), (A1, B2), (B1, B2), (C1, B2), (D1, B2).

It is not necessary to provide all possible values: for example, a binary search technique may be used or a discrete set of parameter combinations may be provided with multiple parameters changing from one set of parameters to the next. For example, the parameter values provided by parameter selector 220 could be limited to (A1, A2), (B1, B2), (C1, A2), (D1, B2).

Each time the parameters are provided to recognition engine, parameter selector 220 provides those parameters or an indicator of the parameters to optimal parameter identifier 222. In addition, recognition engine provides the confidence level of the character recognition to optimal parameter identifier. Optimal parameter identifier 222 records the parameters and the confidence level so that the optimal value for the parameters may be identified by

selecting values of parameters that yield the highest confidence levels.

002030" E2HE950
In one embodiment, optimal parameter identifier identifies the optimal value for a parameter by choosing
5 the value of a parameter that yielded the highest confidence level for the recognition performed by recognition engine 216. In another embodiment, if the confidence level exceeds a threshold, such as 0.95, optimal parameter identifier 222 signals parameter selector 220 to
10 select no further values of that parameter. This way, if an acceptable value is reached, the system 200 does not continue searching for a better one: the acceptable value is used as the optimal value.

In one embodiment, when an optimal value for a
15 particular parameter has been identified, optimal parameter identifier 222 provides the value of the optimal parameter to parameter selector 220 to use in all remaining parameter selections it makes as described above. In another embodiment, a preset value may be used in place of the
20 optimal parameter for use in attempting to identify the optimal value of a different parameter.

When the optimal values of all of the parameters have been identified as described above, optimal parameter

storage 210 using the same parameters as the last segment,
and only signals parameter selector 220 to begin the
process of parameter selection again if the confidence
level of the recognition for the segment falls
5 significantly below the confidence level of the last
recognition or falls below a certain preset value.

Referring now to Figure 3, a method of identifying
optimal parameters for optical object recognition is shown
according to one embodiment of the present invention. The
10 description below uses characters as objects, but an object
may also be noncharacter things as described above. An
digital representation of an image is received and stored
310 at a first resolution, such as a native resolution of a
scanning device as described above. A reduced-resolution
15 version of some or all of the image received in step 310 is
produced and stored 312 as described above. A segment is
selected 314, either from the image received in step 310 or
the version of the image produced in step 312. An initial
set of parameters are selected 316 as described above. An
20 attempt is made to recognize one or more objects such as
characters in the segment from the reduced resolution
version of the image, and the highest confidence level
obtained from the attempt is produced 318. If there are
additional combinations of parameters 320, a different

values of parameters to use for the segment on which the
recognition was attempted in step 340 but acceptable
recognition confidence was not achieved.

00/080"E4HE950